IOWA STATE UNIVERSITY
Digital Repository

Retrospective Theses and Dissertations

Iowa State University Capstones, Theses and Dissertations

1973

Characterization of DNA from enteric bacteria by analysis of thermal transition curves

Melvin Edward Neville
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

 Part of the Microbiology Commons

Recommended Citation
Neville, Melvin Edward, "Characterization of DNA from enteric bacteria by analysis of thermal transition curves " (1973). *Retrospective Theses and Dissertations*. 5109.
https://lib.dr.iastate.edu/rtd/5109

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106

74-9143

NEVILLE, Melvin Edward, 1944-
CHARACTERIZATION OF DNA FROM ENTERIC BACTERIA BY ANALYSIS OF THERMAL TRANSITION CURVES.

Iowa State University, Ph.D., 1973
Bacteriology

University Microfilms, A XEROX Company, Ann Arbor, Michigan

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED.

Characterization of DNA from enteric bacteria by analysis of thermal transition curves

by

Melvin Edward Neville

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major: Bacteriology

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1973

TABLE OF CONTENTS

INTRODUCTION

Although twenty years have passed since Watson and Crick's (1953) work was presented, only recently has the microbiologist begun to use molecular data about deoxyribonucleic acid composition in the taxonomy of bacteria (Gibbons and Buchanan, *in press*). Determinations of characteristics such as DNA base composition now supplement the usual battery of morphological and physiological tests. Perhaps the final step in classifying bacteria will be to determine the sequence of purine and pyrimidine bases in their DNA. Relating this sequence to a map of gene functions would permit a more complete understanding of organisms, and thus make their classification much easier.

But we are not yet close to such a goal. Although there are a few linkage maps, such as those for *Escherichia coli* (Taylor and Trotter, 1972), *Salmonella typhimurium* (Sanderson, 1972), and bacteriophage lambda (Szybalski, 1971), our knowledge of the genetics of other microorganisms is less complete, and detailed linkage maps will not soon be available. Likewise, direct sequence analyses of DNA for organisms are almost non-existent at present; the only determinations have been on the ends of a few viral nucleic acids such as lambda (Wu and Taylor, 1971), coliphage 186 (Padmanabhan and Wu, 1972), and phage T7 (Obel and Freifelder, 1972). The many problems in sequence analysis of DNA from higher organisms have forced investigators to adopt indirect methods such as determining the sequence of particular proteins (Dickerson, 1971) or RNA (*e.g.*, Holley *et al.*, 1965). But these are also enormous research undertakings, and most investigators settle for a less rigorous characterization of the DNA.

One of the most useful properties of DNA from an organism is the overall base composition, usually determined by thermal denaturation (Marmur and Doty, 1962) or ultracentrifugation (Schildkraut, Marmur, and Doty, 1962). More elaborate analyses of DNA include the study of compositional heterogeneity (Miyazawa and Thomas, 1965; Nandi, Wang, and Davidson, 1965; Bellett, 1967; Felsenfeld, 1968) or a study of homologous sequences of nucleotides (McCarthy, 1967; Britten and Kohne, 1968; Brenner and Falkow, 1971). These elaborate analyses are technically difficult, are not readily applicable to large numbers of organisms, and may require elaborate equipment. It would be advantageous to find a simple, reliable technique which would yield more information about DNA.

Three existing spectrophotometric techniques are apparent alternatives. The first technique involves determining DNA homology from renaturation rates of DNA solutions (DeLey, Cattoir, and Reynaerts, 1970; Seidler and Mandel, 1971); results similar to hybridization studies are produced without using radioactive labels. Another technique is to treat thermal transition curves of DNA as cumulative frequency distributions (Krieg and Lockhart, 1970); the statistics describing the distributions are different and distinctive for different organisms. The third technique is spectral analysis, as developed by Felsenfeld and Hirschman (1965, 1966) and expanded by Prouty (1970) and Weiner (1970); one determines the nucleotide composition of successively denaturing portions of a DNA molecule and represents this information as a histogram or "profile" of the DNA from an organism.

The objective of this investigation was to compare and evaluate the spectral data obtained from thermal transitions and from spectral analyses,

to survey a group of enteric strains by these spectral methods, and to relate the data to what is already known about the enterics. The enterics were chosen because extensive genetic, phenetic, and molecular data have been accumulated for them, and because there is a well-developed technique for purifying DNA from gram-negative organisms.

MATERIALS AND METHODS

Organisms Thirty-four enteric bacteria were used in this study. All cultures listed in Table 1 were obtained from the American Type Culture Collection (ATCC), Rockville, Maryland. As noted in the table, most of the strains are taxonomically significant, and many had been characterized by conventional taxonomic tests in previous studies.

Lyophilized cultures were revived in brain heart infusion (BHI) (Difco, 1953). After 15 minutes in BHI, the cultures were transferred to BHI agar plates. Isolated colonies were restreaked onto two plates, and these second transfers served as working cultures. From the working cultures, key tests were performed to confirm the identity of the strains (Edwards and Ewing, 1962; American Type Culture Collection, 1972). Lyophils of each strain were prepared from the working cultures, and these lyophils became my stock cultures.

Data on DNA purified and characterized by other investigators were included in this study for reference purposes. These organisms are listed in Table 2.

Purification of bacterial DNA Cells were grown to stationary phase in BHI, and their DNA was extracted essentially by the Marmur (1961) procedure. Cells were washed twice, suspended in saline-EDTA, and lysed with sodium lauryl sulfate. DNA was digested by ribonuclease (Worthington), and the preparation was deproteinized with a mixture of chloroform-isoamyl alcohol. The DNA was precipitated each time with two volumes of ethanol; in a standard Marmur procedure the final precipitation is with isopropanol. Purified, spooled DNA was stored in 95 % ethyl alcohol at

4 C until needed for analysis.

Immediately before analysis, two or three spools of precipitated DNA were dissolved in approximately 15 ml of phosphate salts buffer (PSB), which contained 0.0085 M KH_2PO_4, 0.0145 M K_2HPO_4, and 0.1000 M NaCl in deionized water, pH 7.0. This DNA solution was dialyzed against four changes of PSB at 4 C (400:1, buffer to DNA solution) over a period of 48 hours. After dialysis, an aliquot of the DNA solution was adjusted spectrophotometrically to a concentration of 20 - 25 µg/ml (optical density 0.4 - 0.5 at 260 nm). Purity of DNA preparations was checked by Lowry, orcinol, and enzymatic methods, and neither protein nor RNA was detected within the limits of sensitivity for these tests. It was concluded that contamination was less than 1.0 %.

Purification of DNA from bacteriophage lambda Bacteriophage lambda was induced from a temperature-sensitive strain of *E. coli* (Wu and Kaiser, 1967). *E. coli* K-12 Strain W3350 (λ *ind*$^-$ $C_{I\ 857}$) was grown at 35 C in 1 % tryptone broth. Cultures were heat-induced at 45 C with gentle swirling for 10 minutes, then shaken at 37 - 39 C until lysis had occurred. From this point the phage DNA was purified as described by Prouty (1970). Low-speed centrifugation removed cell debris, and the lysate was precipitated with a Spinco model L preparative ultracentrifuge. Resuspended phage pellets were purified by "banding" the phage in a cesium chloride gradient in the ultracentrifuge. Once the CsCl had been removed by dialysis, DNA was purified by phenol extraction. The phage suspension was extracted with an equal volume of phenol (50 gm of freshly distilled phenol plus 20 ml 1 M PSB, pH 7.0) by gently swirling the mixture in an

ice bath for 15 minutes. The mixture was centrifuged, and the aqueous layer containing the DNA was removed. This procedure was repeated twice. Residual phenol was removed by washing the aqueous layer three times with absolute ethyl ether, and the residual ether was removed by blowing air over the solution. From this point the lambda DNA solution was treated as previously described for bacterial DNA.

Collecting spectral data for DNA Three ml of DNA solutions were pipetted into 10 mm, matched, teflon-stoppered cuvettes (Thomas). Changes in temperature and absorbance were measured with a Model 2000 multiple-sample recording spectrophotometer (Gilford Instrument Laboratories, Inc., Oberlin, Ohio). Each determination included three DNA samples and a solvent blank. Temperature of the cuvette chamber was regulated by circulating heated ethylene glycol through plates on both sides of the chamber with a manually-controlled Haake Ultrathermostat (Poly Science Corporation, Evanston, Illinois). Temperature was measured through a thermosensor under the cuvette carriage and was recorded on an expanded scale of an ElectroniK 15 Recorder (Honeywell, Philadelphia, Pennsylvania). Absorbance was read from a calibrated Gilford 209 absorbance meter.

The ultraviolet spectrum of the DNA sample at ambient temperature was measured at each of 15 wavelengths, 220 through 290 nm in 5 nm increments. The temperature of the cuvettes was then quickly raised to approximately 3 C below the temperature at which denaturation of any sample was expected to begin. Cuvettes were removed from the heating chamber, opened to relieve pressure and to free air bubbles, and then replaced. Temperature was raised in approximately 0.6 C increments, and

the cuvettes were allowed to equilibrate approximately 20 minutes before readings were made. Equilibrium was assumed when three successive readings of both temperature and absorbance were the same. The absorbance was measured after each temperature increase at one wavelength (260 nm) for an analysis by thermal transition or at 15 wavelengths (220 through 290 nm in 5 nm increments) for a spectral analysis. Denaturation was considered complete when increased temperature produced no further change in the absorbancy at 260 nm.

Characterization of spectral data Thermal transition curves at 260 nm were analyzed by the procedures of Krieg and Lockhart (1970), who treated these curves as cumulative frequency distributions and computed several statistics of each distribution. The T_m was defined as the median temperature, at which 50 % of the total absorbance change had occurred. Mean T (mean temperature) was the temperature at which the average fragment melted and was computed by the formula: mean $T = \Sigma F_i T_i / N$, where F_i was the frequency (*i.e.*, the absorbance change) recorded for each temperature interval, T_i was the midpoint of the temperature interval, and N was the total absorbance change (*i.e.*, ΣF_i). The overall base composition was determined by the formula: G + C = 2.44 (mean T - 81.5 - 16.6 log M) where M is the molar concentration of cations (Mandel and Marmur, 1968). The variance (σ^2) was defined as $\{ N \Sigma F_i(T_i)^2 - \Sigma (F_i T_i)^2 \} / N^2$. The standard deviation (σ) was then derived from the variance. The third standard moment (skewness) and the fourth standard moment (kurtosis) were computed as skewness = $\Sigma\{(F_i/N)(Z_i)^3\}$ and kurtosis = $\Sigma\{(F_i/N)(Z_i)^4\}$, where the standard (Z) value for each temperature was calculated as follows:

$Z_i = (T_i - \text{mean } T) / \sigma$. Finally, the hyperchromic change was determined by the relationship: (final O.D. - initial O.D.) / initial O.D.

Spectral analysis employing the data from 15 wavelengths was performed by the procedure of Prouty (1970). A complete discussion of the rationale, parameters, and formulae of this technique can be found elsewhere, for example, in Prouty (1970) or Felsenfeld (1968); briefly, the technique determines the overall base composition, the amount of denaturation, the fraction of total DNA which has been denatured during each temperature interval, and the base composition of each such fraction. Both two-term (4 wavelengths: 250, 260, 270 and 280 nm) and three-term analyses (all 15 wavelengths) are possible by this technique, and theoretically both should give similar results.

RESULTS

Comparison of spectral techniques The first goal of this study was to determine which technique involving denaturation of DNA could better detect relatedness among organisms. Typical data from transition curves for 10 of the 15 wavelengths measured in a spectral analysis are shown in Table 3. These characteristics were essentially the same for most wavelengths except at the extremes, 220 and 290 nm. The absorbancy change was greatest at 260 and 265 nm, emphasizing that this would be the most sensitive part of the spectrum in which to determine transitions if only a single wavelength were employed.

A typical denaturation profile determined by spectral analysis is shown in Figure 1 for DNA of *E. coli* ATCC 9637. The profile is similar to the ones presented by Prouty (1970) and Weiner (1970) for the same strain. The peaks represent values obtained when the results of two-term and three-term analyses were combined; the hash marks represent values obtained when either two-term or three-term results were averaged alone. Obviously the results of two-term and three-term analysis are in poor agreement. Furthermore, there was a great range in the individual values from which these means were derived, even when values were for replicate samples from the same determination. Standard errors were calculated for each peak, and in many cases these standard errors were more than 20 % of the peak itself. Thus, despite the fact that spectral analysis appears to have the advantage of determining many parameters of a DNA solution, the technique yields results that are poorly reproducible.

Analysis by thermal transition was then checked for reproducibility.

Table 4 shows typical data for DNA purified from three different batch cultures of *E. coli* ATCC 9637. Thermal transitions were determined using material from the three purifications in different combinations. For example, one determination used three samples from the same purification, and another determination used a sample from each of the three purifications. By analyzing these and other, similar determinations, it was found that for DNA from a particular strain the greatest variation in the thermal transition occurred between determinations, whereas the three samples included in any given determination produced essentially identical results. Therefore, the most critical comparisons between samples could be made by performing thermal denaturations in the same determination. To obtain representative variation, replicate samples for each strain were analyzed in a series of different determinations.

Under these conditions, the means of three determinations for DNA from a given strain produced statistics that were characteristic and reproducible. The means did not change appreciably if additional determinations were performed, although standard errors might be decreased. In any event the standard errors were quite small. Except for the overall percent G + C in four cases, for the standard deviation in one case, and for kurtosis in three cases, the standard errors for all characteristics of DNA from each of the 39 strains studied fell within the limits shown in Table 5. These values are only slightly better than those reported as standard errors by Krieg and Lockhart (1970), emphasizing that these variations are probably inherent in the technique; this variation might result from differences in the fragmentation of purified DNA or from limitations

of the instruments.

Effects of various parameters on thermal transitions Because analysis by thermal transitions seemed a more reproducible method than spectral analysis, thermal transitions were used to compare bacterial DNA in this study. Before making such comparisons, it was advisable to investigate the technique for sources of variation. Using *E. coli* ATCC 11775, I first varied the purification procedures by eliminating the precipitation of the DNA with ethanol after each deproteinization. Within the limits of accuracy shown in Table 5, this made no difference in the characteristics of the transition curves. However, eliminating the precipitation steps greatly reduced the time required to purify the DNA. The apparent yield of DNA was also increased. A second modification was to subject DNA purified by my modified Marmur procedure to additional phenol extractions. Purified DNA of *E. coli* ATCC 9637 was dissolved in PSB at an approximate concentration of 25 µg/ml. After the DNA had completely dissolved, the sample was divided into two aliquots. One aliquot was further purified by the phenol method described earlier for bacteriophage lambda. DNA prepared in this manner had identical characteristics to the other aliquot prepared without the phenol steps. Another slight modification was to use frozen cells which had been stored for two weeks instead of cells that had just reached stationary phase. Two liters of cells were grown to stationary phase, washed twice in saline-EDTA, resuspended in saline-EDTA, and divided into two aliquots. One aliquot was frozen, and from the other aliquot DNA was purified as described earlier. Two weeks later the frozen cells were thawed, and DNA was purified from them. There

were no differences in the characteristics of these two preparations, so frozen cells could be used equally well. It was concluded that the reproducibility of this technique is not particularly sensitive to minor differences in the extraction procedure.

Even though protein and RNA assays on the purified DNA samples did not detect significant contamination, one source of variation in thermal transition curves might have been that contaminating materials survived the purification procedure. Commercially purified yeast RNA (Sigma) and ribonuclease (Worthington) at final concentrations of 2.0 and 20.0 µg/ml were added to a DNA solution (20 µg/ml) of *E. coli* ATCC 11775, and thermal transitions were determined. As can be seen in Table 6, adding ribonuclease to the purified DNA did not appreciably affect the transition curves; all the values were within the standard errors considered to be acceptable. Likewise, when RNA was added, there were no significant changes; the only difference was in the hyperchromic change, which was less because the RNA also absorbs well at 260 nm. It is not possible to generalize and say that no contaminants would have any effect on the thermal transition curves, but it appeared that at least the most likely contaminants in a DNA purification were not a significant source of error.

Next, the different factors involved in measuring the thermal transition were investigated. Because our laboratory had been using phosphate salts buffer as the solvent for determinations, I needed to show what effects different solvents might have on the characteristics. Four spools of purified DNA from *E. coli* ATCC 9637 were dissolved in a mixture of PSB and standard saline citrate (SSC). After the DNA had completely dissolved,

the solution was divided into two parts, and each was dialyzed against only one of these buffers according to the same dialysis schedule. As can be seen in Table 7, there were no differences in the characteristics of DNA determined in PSB and SSC. The statistics were the same except for the denaturation temperatures; however, denaturation temperatures are a function of the cation concentration (Schildkraut and Lifson, 1965), and when mean percent G + C was calculated by the appropriate formulae, the overall base compositions were the same.

To determine if small variations in pH could affect the transition curves, purified DNA was dialyzed against PSB in the pH range 6.0 through 8.0. Six spools of purified DNA from *E. coli* ATCC 4157 were dissolved in 75 ml of PSB at an approximate concentration of 24 µg/ml. After the DNA had completely dissolved, the solution was divided into seven aliquots, and each aliquot was dialyzed against PSB at only one of these pH's. The dialysis solutions were prepared by mixing two stock solutions of the phosphate salts (0.1375 M Na^+ + K^+) until the desired pH was attained. Table 8 shows that from pH 6.0 through 8.0, a change in pH had little effect on the transition curves. Except for a slight lowering of the denaturation temperatures at pH 6.0, the characteristics were essentially unchanged over this range.

One variable that might contribute to variation among thermal transitions was the temperature intervals between successive absorbance measurements; these necessarily were different from one determination to the next. Several transitions were performed with *E. coli* ATCC 11775 with measurements at intervals of approximately 0.25 C. This made it

possible to select data for computations (for example, to use the absorbance determined at every other temperature) and thus to determine the effects these intervals might have on the characteristics of the transition curves. The results of varying the intervals by manipulating data collected at the smallest increments physically possible are shown in Table 9. The characteristics of the thermal transitions were not greatly altered; all values were well within the acceptable standard errors. The only value affected was the T_m, again suggesting that the mean T may be a more reliable statistic for computing overall base composition (Krieg and Lockhart, 1970). Therefore, it seemed that differences in the temperature intervals were not an important source of variation.

Characteristics of the thermal transitions It became quite obvious that once the limits of reproducibility of the technique were determined, such factors as purification procedures, solvents, pH and temperature intervals did not greatly affect the reproducibility of the technique. The characteristics of the 34 enteric strains were determined, and their means plus those of the five reference strains and bacteriophage lambda are listed in Table 10. Superficially, the data were as expected. As noted earlier, there were only a few values that did not fall within the predicted standard errors. All the denaturation temperatures were within 1 C of those in compiled lists (DeLey, 1970). Values of standard deviation were in a relatively narrow range (2.4 - 3.3), values of skewness were all negative, and values of kurtosis had the greatest range (2.4 - 6.1); these data are consistent with those presented by other investigators (Krieg, 1968; Prouty, 1970; Weiner, 1970), as shown in Table 11. It is apparent

that the characteristics are readily reproducible by different investigators.

The standard deviation is a measure of dispersion about the mean; a greater value for standard deviation indicates greater intramolecular heterogeneity of the DNA. Skewness is a measure of the asymmetry of the transition; a value of 0.0 indicates no skewness, while a value less than zero indicates a skewness toward lower temperatures. Kurtosis is a measure of sharpness of the transition; a normal distribution has a value of 3.0 for kurtosis, a value less than 3.0 indicates a flattened curve, and a value greater than 3.0 indicates a peaked curve. A large value for kurtosis (greater than 3.0) indicates the presence of many homogeneous DNA fragments.

Comparisons between characteristics At this point it became necessary to determine whether the characteristics were independent of one another - that is, was each value a measure of a different property of the thermal transition, or were they all measuring essentially the same thing? To reduce all the values to a common basis, and thus to facilitate comparisons, a "difference" table was constructed for each characteristic. The value of any characteristic was considered to be "different" between any pair of organisms if they differed from one another by as much as the standard error shown in Table 5 (*i.e.*, if the difference between them was greater than the experimental error in determining the value in question). If the values differed by twice the standard error, two differences were recorded between the two organisms; if the values differed by three times the standard errors, three differenences were recorded, *etc.* Thus, for

each set of values, a table was constructed which contained a difference value for each possible comparison between pairs of strains (741 values, for all comparisons among the 39 strains in this study). Difference tables for the various values were then compared by computing the Pearson product-moment correlation coefficients between them, using an appropriate FORTRAN program (Lockhart, 1967). Perfectly correlated tables would yield a coefficient of 1.0, while a value of 0.0 would indicate that no correlation whatsoever existed between two tables. The minimal value for a correlation coefficient that can be considered significantly different from zero for various number of values compared may be found in statistical tables (*e.g.*, Snedecor and Cochran, 1967).

When data from all 39 bacterial strains were compared (Table 12), the characteristics were correlated with one another at the 1 % significance level except for the mean percent G + C compared with kurtosis. Although this suggested that the characteristics were providing only slightly more information than was already known from G + C alone, the comparisons may have been biased by the fact that most of the organisms studied were closely related to one another and contained DNA closely similar in base content and the characteristics of thermal transition curves. When 10 unrelated strains with a wide range of mean percent G + C were selected, and correlation coefficients were calculated between the characteristics of their transition curves (45 pairwise comparisons), the only significant correlation was between the mean percent G + C and skewness (Table 12). Probably, therefore, most of the characteristics were independent of one another; each of the characteristics apparently measured something unique

about the DNA. Even the "significant" correlation between mean percent G + C and skewness was not high, and it was decided to include skewness in subsequent computations also.

Preparation of diagrams The characteristics of their thermal transitions were then used to sort strains into groups by the methods of numerical taxonomy (Lockhart and Liston, 1970). First, the tables of difference values for standard deviation, skewness, and kurtosis were combined to produce a table which included all differences except those in total base content (mean percent G + C), and the strains were sorted by a "single-linkage" procedure in which individuals were joined to groups at the level of their fewest differences with any other member of the group. The relationships found among strains by this technique were superimposed on a diagram in which the strains were arranged on an axis of mean percent G + C. Figures 2 and 3, therefore, represent diagrams in which the strains are separated primarily by the base content of their DNA, and then are additionaly separated by other differences in the characteristics of their thermal transitions. This diagram substantiates the findings of Krieg and Lockhart (1970) that the values of a thermal transition are distinctive for a given organism, and that more closely related organisms have similar characteristics.

The enterics, grouped in a cluster between 49 and 60 % G + C, joined at a level of four differences. Closer inspection of this group, as seen in Figure 3, showed that the eleven strains of the genus *Escherichia* formed a cluster with *Citrobacter freundii* ATCC 6750, *Proteus morganii*, two strains of *Erwinia carotovora*, and *Salmonella typhimurium*. It is interesting to

note that *Neisseria sicca* was also found in this group. Two strains of *Erwinia*, *E. amylovora* and *E. nimipressularis*, formed a cluster with the enterobacters, but *Erwinia carnegieana* grouped with the klebsiellas and serratias.

Relationship of findings to other taxonomic studies As was shown in Table 1, many of these organisms had been included in earlier taxonomic studies in which the organisms were sorted into numerical groupings on the basis of conventional diagnostic tests. The method of Lockhart (1967) was used to compute correlation coefficients comparing the relative positions of organisms in those diagrams of Lockhart and Koenig (1965), Focht and Lockhart (1965), and Krieg and Lockhart (1966) with their positions in the diagram based on DNA characteristics just discussed. The resulting correlations, based on 21, 55, and 120 pairwise comparisons respectively, are shown in Table 13. The grouping by DNA characteristics was significantly similar to those groupings derived from numerical taxonomy based on conventional diagnostic tests.

TABLES AND FIGURES

Table 1. Organisms used in this investigation

Code Number	Organism	ATCC [a] Number	Nomenclatural status [b]	Phenetic study [c]
1	*Escherichia coli* (Strain W)	9637		Kr
2	*E. coli* (Escherich strain)	4157		
3	*E. coli*	11775	N	Fo, Kr, Lo
4	*E. coli* (Wild type B)	23848		
5	*E. coli* (Strain C)	13706		
6	*E. coli* (Wild type K-12)	23716		
7	*E. coli* (K-12 Hfr $(\lambda)^-$)	23739		
8	*E. coli* (K-12 Hfr $(\lambda)^+$)	23740		
9	*E. coli* (K-12 Hfr met$^-(\lambda)^+$)	23741		
10	*E. coli*	23742		
11	*E. aurescens*	12814	R	Kr
12	*Enterobacter aerogenes*	13048	N	Kr
13	*E. cloacae*	13047	N	Fo, Kr, Lo
14	*E. liquefaciens*	14460	T	Fo, Kr, Lo
15	*E. liquefaciens*	14461	R	
16	*Serratia marcescens*	13880	N	Fo, Kr, Lo
17	*S. marcescens* subsp *kiliensis*	8101	R	Fo, Kr
18	*Klebsiella pneumoniae*	13882		
19	*K. pneumoniae*	13883	N	Kr
20	*K. rhinoscleromatis*	13884	N	Kr
21	*K. edwardsii* subsp *edwardsii*	13886		
22	*K. edwardsii* subsp *atlantae*	13887		
23	*Proteus morganii*	25830	R	
24	*P. vulgaris*	13315	N	Fo, Kr, Lo
25	*Erwinia carotovora*	495	R	Fo, Kr, Lo
26	*E. carotovora*	15713	N	
27	*E. amylovora*	15580	N	
28	*E. nimipressularis*	9912	R	Lo
29	*E. cassavae*	23372	T	
30	*E. carnegieana*	13452	R	
31	*Arizona arizonae*	13314	N	Kr
32	*Citrobacter freundii*	8090	N	Kr
33	*C. freundii*	6750		
34	*Paracolobactrum aerogenoides*	11604		Kr

[a] American Type Culture Collection, Rockville, Maryland.

[b] N = Neotype strain (American Type Culture Collection, 1972).
R = Representative strain (American Type Culture Collection, 1972).
T = Type strain (American Type Culture Collection, 1972).

[c] Fo = Focht and Lockhart (1965).
Kr = Krieg and Lockhart (1966).
Lo = Lockhart and Koenig (1965).

Table 2. Organisms compared in this study but characterized by other investigators

Code Number	Organism	Investigator
35	*Neisseria sicca*	Weiner (1970)
36	*Pseudomonas aeruginosa*	Prouty (1970)
37	*Staphylococcus aureus*	Prouty (1970)
38	*Bacillus subtilis* (Strain W23)	Needleman [a]
39	*Salmonella typhimurium*	Krieg (1968)

[a] David S. Needleman, 1972, *Personal communication*.

Table 3. The characteristics of thermal transition curves for DNA (22 µg/ml) purified from *E. coli* ATCC 9637 as measured at different wavelengths

Wavelength (nm)	Denaturation temperatures Mean	T_m	Mean percent G + C	Standard deviation	Skewness	Kurtosis	Hyperchromic change (1000 x O.D.)
220	88.2	88.7	51.1	2.6	- 0.9	4.0	98
230	88.3	88.6	51.4	2.4	- 1.0	4.0	76
240	88.4	88.7	51.6	2.3	- 0.8	4.0	113
250	88.3	88.7	51.4	2.4	- 0.8	3.7	146
255	88.1	88.6	50.9	2.5	- 1.0	3.9	162
260	88.1	88.6	50.9	2.6	- 0.9	3.9	170
265	88.2	88.6	51.1	2.5	- 0.8	3.6	172
270	88.1	88.6	50.9	2.5	- 1.0	4.0	164
280	88.3	89.1	51.4	2.5	- 1.0	3.9	102
290	87.7	88.6	49.9	3.0	- 1.0	3.2	31

Table 4. Characteristics of fourteen thermal transition curves at 260 nm for DNA purified from *E. coli* ATCC 9637

	Denaturation temperatures		Mean percent G + C	Standard deviation	Skewness	Kurtosis	Percent hyperchromic change
	Mean	T_m					
Means	88.4	88.2	51.7	2.6	- 0.9	3.9	37.1
Standard errors	0.2	0.2	0.4	0.1	0.1	0.2	1.3

Table 5. Reproducibility of characteristics of thermal transitions for 39 bacterial strains

Characteristic	Standard Error
Mean percent G + C	0.6 %
Standard deviation	0.2 C
Skewness	0.2
Kurtosis	0.4

Table 6. Effects of contaminants on the characteristics of thermal transition curves for DNA (20 µg/ml) from *E. coli* ATCC 11775. A control and two concentrations for each contaminant are from the same determination

Contaminant (µg/ml)	Mean percent G + C	Standard deviation	Skewness	Kurtosis	Percent hyperchromic change
Ribonuclease					
None	51.7	2.4	- 0.9	4.2	33.7
2	51.7	2.5	- 1.0	4.2	37.6
20	51.9	2.6	- 1.0	4.1	36.1
Ribonucleic acid					
None	51.5	2.6	- 1.1	4.3	37.0
2	51.8	2.6	- 1.1	4.2	33.2
20	51.8	2.7	- 1.1	4.2	18.4

Table 7. Effects of solvents on the characteristics of thermal transition curves for DNA from *E. coli* ATCC 9637. Each set of values represents a sample from the same determination

Solvent	Denaturation temperatures Mean	T_m	Mean percent G + C	Standard deviation	Skewness	Kurtosis	Percent hyperchromic change
PSB [a]	88.4	89.0	51.6	2.5	- 1.1	4.3	36.4
SSC [b]	90.6	91.2	51.6	2.6	- 1.1	4.3	34.4

[a] PSB = Phosphate salts buffer, pH 7.0 (0.1375 M $Na^+ + K^+$).

[b] SSC = Standard saline citrate, pH 7.0 (0.185 M Na^+).

Table 8. Effects of pH on the characteristics of thermal transition curves for DNA from *E. coli* ATCC 4157. Each value represents an average of two or more determinations

pH	Denaturation temperatures Mean	T_m	Standard deviation	Skewness	Kurtosis	Hyperchromic change (1000 x O.D.)
6.0	86.9	87.3	2.5	- 0.7	3.6	160
6.5	87.7	88.2	2.6	- 0.9	3.7	158
6.8	87.7	88.3	2.7	- 0.9	3.8	163
7.0 [a]	88.1	88.6	2.6	- 1.0	4.0	160
7.0 [b]	88.0	88.5	2.6	- 1.0	3.9	158
7.2	88.0	88.4	2.5	- 1.0	3.8	159
7.5	88.0	88.4	2.5	- 1.0	3.8	159
8.0	87.5	88.1	2.6	- 1.0	3.9	163

[a] Phosphate salts buffer made according to formula.

[b] Phosphate salts buffer adjusted to pH 7.0 as by this experiment.

Table 9. Effects of temperature intervals on the apparent characteristics of a thermal transition curve at 260 nm for DNA from *E. coli* ATCC 11775

Average increment (C)	Denaturation temperatures Mean	T_m	Standard deviation	Skewness	Kurtosis
0.25	88.6	89.0	2.4	- 0.9	3.6
0.5	88.7	88.9	2.5	- 0.9	3.6
1.0	88.7	89.2	2.4	- 0.9	3.6
1.5	88.7	89.4	2.5	- 0.8	3.5
2.5	88.7	90.2	2.5	- 0.8	3.5

Table 10. Mean characteristics of thermal transition curves at 260 nm for DNA purified from each of the organisms in this study, and the number of determinations from which each mean was computed

Code Number	Organism	Mean percent G + C	Denaturation temperatures Mean	T_m	Standard deviation	Skewness	Kurtosis	No. of determ.
1	*Escherichia coli*	51.7	88.4	88.8	2.6	- 0.9	3.9	14
2	*E. coli*	50.9	88.1	88.6	2.8	- 1.0	3.9	7
3	*E. coli*	51.8	88.4	88.9	2.6	- 0.9	4.0	9
4	*E. coli*	51.6	88.3	88.8	2.7	- 1.1	4.2	5
5	*E. coli*	52.1	88.6	89.1	2.6	- 0.9	4.0	3
6	*E. coli*	51.6	88.4	88.8	2.6	- 1.0	4.0	5
7	*E. coli*	51.6	88.3	88.8	2.6	- 0.9	3.8	3
8	*E. coli*	52.1	88.5	89.2	2.7	- 0.9	3.7	3
9	*E. coli*	51.8	88.4	88.9	2.5	- 0.8	3.7	3
10	*E. coli*	51.8	88.4	89.0	2.8	- 1.1	4.4	3
11	*E. aurescens*	50.8	88.0	88.6	3.0	- 1.3	5.1	3
12	*Enterobacter aerogenes*	55.4	89.9	90.2	2.8	- 0.8	3.6	5
13	*E. cloacae*	54.3	89.4	89.9	2.7	- 0.7	3.5	3
14	*E. liquefaciens*	53.4	89.1	89.6	2.4	- 1.2	4.8	3

Table 10. (Continued)

Code Number	Organism	Mean percent G + C	Denaturation temperatures Mean	T_m	Standard deviation	Skewness	Kurtosis	No. of determ.
15	*Enterobacter liquefaciens*	55.0	89.7	90.1	2.5	- 1.3	5.4	3
16	*Serratia marcescens*	59.6	91.6	92.2	3.1	- 1.1	4.0	5
17	*S. marcescens* subsp *kiliensis*	59.0	91.4	92.1	3.2	- 1.1	3.9	3
18	*Klebsiella pneumoniae*	58.4	91.1	91.6	2.9	- 1.0	4.3	3
19	*K. pneumoniae*	57.2	90.7	91.1	3.1	- 0.9	3.6	3
20	*K. rhinoscleromatis*	57.3	90.7	91.2	2.9	- 0.8	3.6	3
21	*K. edwardsii* subsp *edwardsii*	59.6	91.6	92.0	3.0	- 0.8	3.8	3
22	*K. edwardsii* subsp *atlantae*	56.9	90.5	90.8	3.0	- 0.8	3.7	3
23	*Proteus morganii*	51.4	88.2	89.0	3.3	- 1.4	5.0	3
24	*P. vulgaris*	38.9	83.1	83.5	2.9	- 0.1	4.2	3
25	*Erwinia carotovora*	51.1	88.2	88.8	3.1	- 1.1	4.4	6
26	*E. carotovora*	51.4	88.3	88.9	3.2	- 1.1	4.3	6
27	*E. amylovora*	54.2	89.4	89.8	3.0	- 0.6	3.6	3
28	*E. nimipressularis*	54.7	89.6	90.1	2.9	- 1.0	4.3	6

Table 10. (Continued)

Code Number	Organism	Mean percent G + C	Denaturation temperatures Mean	T_m	Standard deviation	Skewness	Kurtosis	No. of determ.
29	*Erwinia cassavae*	54.8	89.7	90.2	3.2	- 1.5	6.1	3
30	*E. carnegieana*	57.6	90.8	91.5	3.2	- 0.9	3.6	3
31	*Arizona arizonae*	53.0	88.9	89.4	3.1	- 0.9	3.6	3
32	*Citrobacter freundii*	53.2	89.0	89.4	2.6	- 1.1	4.4	3
33	*C. freundii*	51.3	88.2	89.0	3.2	- 1.2	4.5	3
34	*Paracolobactrum aerogenoides*	48.8	87.2	87.7	2.3	- 0.7	3.7	3
35	*Neisseria sicca*	52.4	88.7	89.3	3.1	- 0.6	3.0	6
36	*Pseudomonas aeruginosa*	62.4	92.7	93.1	2.4	- 0.7	3.3	3
37	*Staphylococcus aureus*	31.6	80.1	80.2	2.0	- 0.1	3.0	2
38	*Bacillus subtilis*	45.4	85.7	86.1	2.4	- 0.6	3.2	6
39	*Salmonella typhimurium* [a]	51.9	90.6	91.0	2.8	- 0.9	4.1	3
40	Bacteriophage lambda	49.9	87.6	88.2	3.2	- 0.5	2.4	10

[a] Determined in standard saline citrate.

Table 11. Characteristics of thermal transition curves at 260 nm for strains included in other investigations

Code Number	Organism	Investigation	Mean percent G + C	Standard deviation	Skewness	Kurtosis
1	*E. coli*	This study	51.7	2.6	- 0.9	3.9
		Prouty (1970)	51.0	2.7	- 0.8	3.6
		Weiner (1970)	52.4	2.7	- 0.8	3.6
3	*E. coli*	This study	51.8	2.6	- 0.9	4.0
		Krieg (1968)	52.1	2.7	- 1.0	4.4
13	*E. cloacae*	This study	54.3	2.7	- 0.7	3.5
		Krieg (1968)	55.7	2.7	- 0.8	3.6
16	*S. marcescens*	This study	59.6	3.1	- 1.1	4.0
		Krieg (1968)	58.6	3.2	- 1.0	4.0
24	*P. vulgaris*	This study	38.9	2.9	- 0.1	4.2
		Krieg (1968)	39.3	2.7	- 0.6	3.3
40	Phage lambda	This study	49.9	3.2	- 0.5	2.4
		Prouty (1970)	51.4	3.3	- 0.4	2.4
		Weiner (1970)	51.4	3.3	- 0.4	2.4

Table 12. Correlation coefficients between characteristics of thermal transitions computed for DNA from 39 strains and from 10 strains[a]

Characteristics compared	Correlation coefficients for DNA from 39 strains	Correlation coefficients for DNA from 10 strains
G + C versus Standard deviation	0.330 **	0.299
G + C versus Skewness	0.539 **	0.651 **
G + C versus Kurtosis	0.067	- 0.013
Standard deviation versus Skewness	0.252 **	0.210
Standard deviation versus Kurtosis	0.184 **	0.199
Skewness versus Kurtosis	0.621 **	0.217

[a] 10 strains that have a wide range of G + C, namely *P. aeruginosa*, 36; *S. marcescens*, 16; *E. cloacae*, 13; *N. sicca*, 35; *E. coli*, 3; *S. typhimurium*, 39; *P. aerogenoides*, 34; *B. subtilis*, 35; *S. aureus*, 37; and *P. vulgaris*, 24.

** Correlation significant at 1 % level; 1 % significance levels for comparing 39 and 10 strains are 0.115 and 0.372, respectively.

Table 13. Correlation coefficients between relationships based on characteristics of DNA and those based on conventional characteristics of organisms

Conventional study	Correlation between conventional and DNA relationships	Significant correlation for this number of comparisons (1 % level)
Lockhart and Koenig (1965)	0.991	0.537
Focht and Lockhart (1965)	0.687	0.372
Krieg and Lockhart (1966)	0.295	0.254

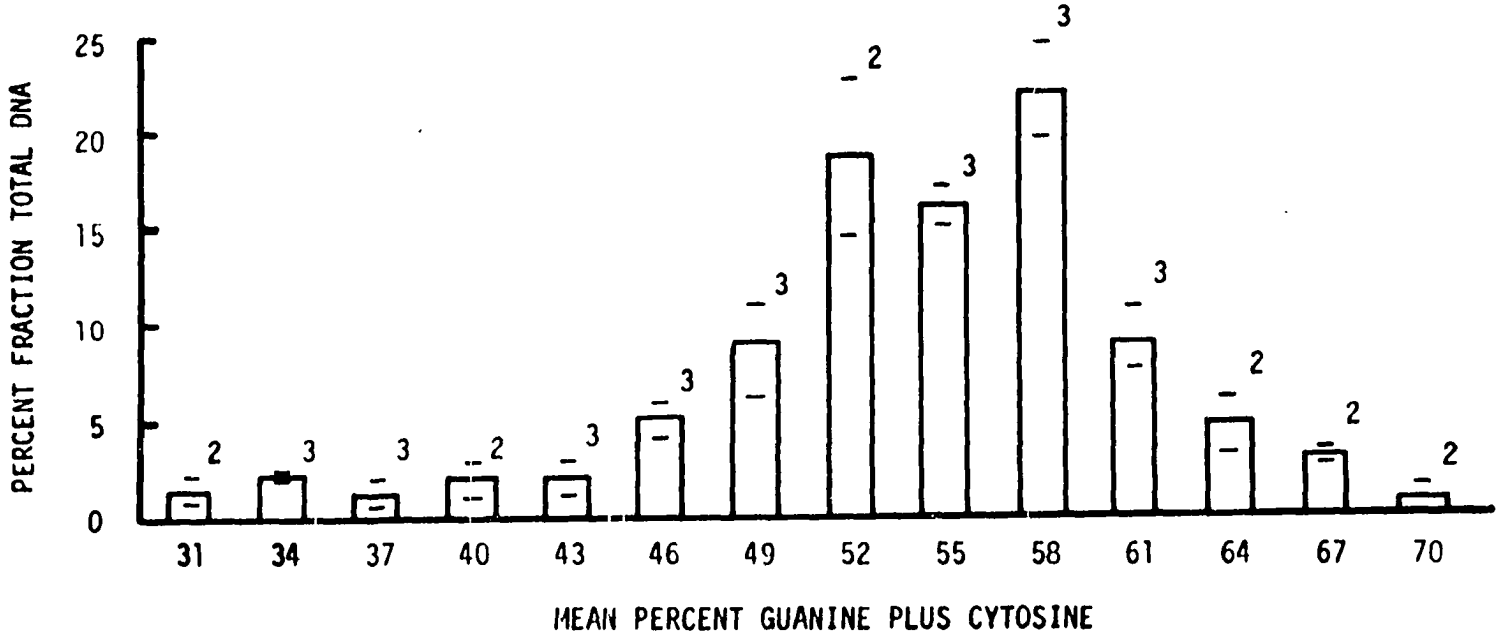


Figure 1. Profile of DNA purified from *E. coli* ATCC 9637 as determined by spectral analysis. Histogram is an average of all data from seven determinations; separate averages of two-term (2) and three-term (3) analyses alone are indicated as hash marks

Figure 2. Diagram of the numerical sortings for strains in this study. Strains were located on an axis of mean percent G + C and separated by the number of differences in the characteristics of their thermal transition curves. The numbers preceding the names of the organisms are the code numbers

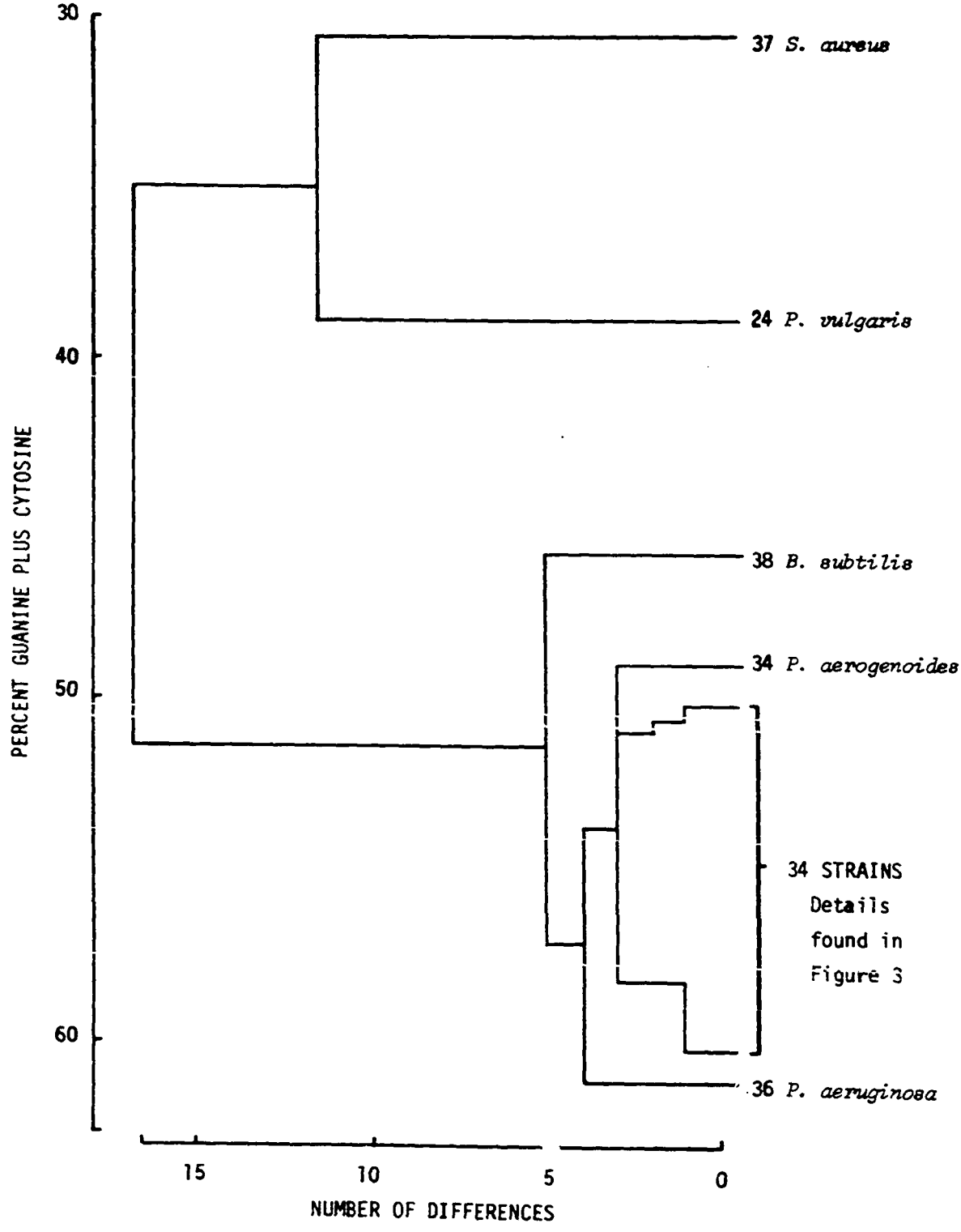
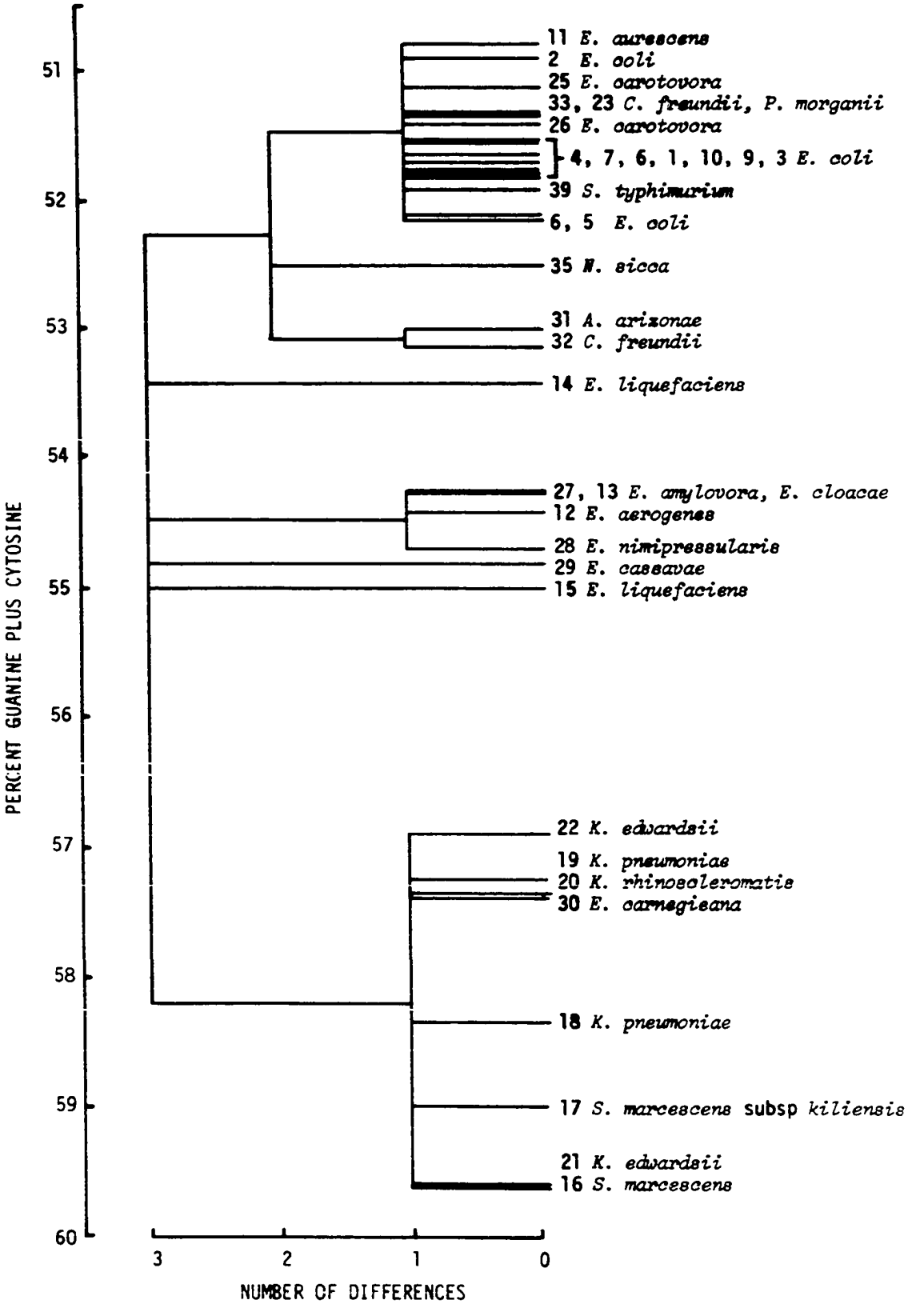

30
40
50
60
PERCENT GUANINE PLUS CYTOSINE
37 S. aureus
24 P. vulgaris
38 B. subtilis
34 P. aerogenoides
34 STRAINS
Details
found in
Figure 3
36 P. aeruginosa
15
10
5
0
NUMBER OF DIFFERENCES

Figure 3. Diagram showing a detailed section of Figure 2. Groups were formed as described for Figure 2. The numbers preceding the names of the organisms are the code numbers


11 E. aurescens
2 E. coli
25 E. carotovora
33, 23 C. freundii, P. morganii
26 E. carotovora
4, 7, 6, 1, 10, 9, 3 E. coli
39 S. typhimurium
6, 5 E. coli
35 N. sicca
31 A. arizonae
32 C. freundii
14 E. liquefaciens
27, 13 E. amylovora, E. cloacae
12 E. aerogenes
28 E. nimipressularis
29 E. casseavae
15 E. liquefaciens
22 K. edwardsii
19 K. pneumoniae
20 K. rhinoscleromatis
30 E. carnegieana
18 K. pneumoniae
17 S. marcescens subsp kiliensis
21 K. edwardsii
16 S. marcescens
51
52
53
54
55
56
57
58
59
60
PERCENT GUANINE PLUS CYTOSINE
3
2
1
0
NUMBER OF DIFFERENCES

DISCUSSION

Using the technique of spectral analysis, Hirschman *et al.* (1967) and Votavova, Sponar, and Sormova (1968) were able to confirm intramolecular heterogeneity in the DNA of bacteriophage lambda and some bacteria. Prouty (1970) and Weiner (1970) presented the results of spectral analysis as DNA profiles in an attempt to facilitate comparisons between DNA from different organisms. However, as shown in this study, the reproducibility of the method is very poor, and these profiles may be no more than artifacts. Spectral analysis is a time-consuming procedure that is not easily comprehended - purified DNA is heat denatured over periods of 12 - 15 hours, and spectral data at several different wavelengths are processed (out of necessity by a computer) via complex mathematical equations which depend upon a number of experimentally-determined constants (Felsenfeld and Hirschman, 1965). Recently, Ulitzur (1972) suggested that the DNA samples from which Felsenfeld and Hirschman derived their constants were contaminated by protein or other impurities, which gives one even less confidence in the technique.

On the other hand, I have shown that the analysis of thermal transitions as measured at a single wavelength can be reproduced within rather stringent limits, and that many of the possible sources of experimental error have only insignificant effects on this reproducibility. Determination of the overall base composition of microbial DNA by thermal transition (Marmur and Doty, 1962) can be reproduced routinely within 1.0 % if the precautions of Mandel and Marmur (1968) are followed, and this reproducibility of the denaturation curve for a particular DNA solution suggests

that additional interpretations of the curve should be possible.

The width of the curve (*i.e.*, the temperature range of the thermal transition) is nearly independent of the size of the DNA fragments in the range of 1 - 25 x 10^6 daltons (Krieg, 1968; DeLey, 1969); however, the width narrows when the size of the fragments increases to 1.3 x 10^8 daltons, and the width increases when the molecular weight is below 6 x 10^5 daltons (Crothers, Kallenbach, and Zimm, 1965). Since it is generally accepted that the molecular weight of DNA purified by the procedure of Marmur (1961) is approximately 10^7 daltons, molecular weight should not be a significant source of variation in the width of transition curves. If a denaturation curve is viewed as the melting of a population of DNA fragments, its width is a measure of the dispersion about the mean (*i.e.*, the standard deviation) of the distribution. The standard deviation, therefore, is a measure of intramolecular heterogeneity of the DNA. Several methods have been suggested for calculating the standard deviation (Knittel *et al.*, 1968; DeLey, 1969; Krieg and Lockhart, 1970), and although each method of calculation may yield a different value for DNA from a given organism, the values seem to be relative. For example, *Serratia marcescens* always has one of the highest values for standard deviation, regardless of the method used to compute it.

Denaturation curves are not usually symmetrical; they are skewed to the left, or toward lower temperatures (DeLey and VanMuylem, 1963; Knittel *et al.*, 1968; Krieg and Lockhart, 1970). I found, as Krieg and Lockhart (1970) had reported earlier, that this skewness tends to be most pronounced in DNA that is high in G + C content (*i.e.*, has relatively few A-T pairs).

Although several theories have been suggested to explain persistent skewness (*e.g.*, difference in hyperchromicity between A-T and G-C pairs or fragility of A-T rich regions), Crothers (1968) and DeLey (1969) believe that the asymmetry is due to the existence of several A-T rich regions.

Another feature of the thermal transition curve is its sharpness or peakedness. Differences in slopes of the curves had been recognized by several investigators, and Krieg and Lockhart (1970) suggested a way to calculate this characteristic. The kurtosis is a measure of the homogeneity of the base composition of the fragments.

After determining these characteristics for DNA from several bacterial species, Krieg and Lockhart (1970) concluded that the denaturation curves are characteristic and constant for the DNA from a given organism, and that they are different and distinctive for different organisms. The results presented here substantiate their findings. It has been shown also that DNA from closely related organisms has characteristics that are more alike than DNA from organisms more distantly related. These apparent similarities and differences are a reflection of the nucleotide sequences in the DNA. Groupings of organisms based on characteristics of their DNA are quite similar to those found by numerical studies based on phenetic characters. There is a large amount of correlation between DNA groupings and the phenetic groupings of Focht and Lockhart (1965) involving diverse bacteria, and between DNA groupings and those of Lockhart and Koenig (1965) involving mainly the erwinias. But for the study of Krieg and Lockhart (1966), involving a spectrum of closely related enterics, the correlation drops drastically (although it is still significant).

There is good agreement between these DNA groupings and the way most of the organisms will be classified in the Eighth Edition of Bergey's Manual (Gibbons and Buchanan, *in press*). The distantly related organisms such as *B. subtilis*, *S. aureus*, and *P. aeruginosa* are separated from one another, as would be expected. On the other hand, *N. sicca* is grouped, mistakenly, with the enterics. The shape of the transition curve of DNA from *N. sicca* was not sufficiently different from those of the enterics to prevent its being grouped with the latter as a result of its coincidental similarity in overall base composition. As in other numerical sortings, we are undoubtedly dealing with a multiply-dimensioned array of groupings, and when the array is condensed into two dimensions, such misrepresentations can be expected occasionally. Otherwise, the DNA groupings of enterics are similar to conventional classifications (Cowan, *in press*). The *Escherichieae* formed the largest cluster in this study, including many of the organisms which will make up the *Escherichieae* in Bergey's. The *Escherichieae* will consist of the genera *Escherichia*, *Edwardsiella*, *Citrobacter*, *Salmonella*, and *Shigella*. The *Klebsielleae* formed a cluster in this study, while in Bergey's the klebsiellas, serratias, enterobacters, and hafnias will make up the *Klebsielleae*. *P. vulgaris* is set apart from the enterics not only by mean percent G + C, but by other characteristics of the denaturation curve, another point in agreement with Cowan (*in press*). It can be concluded that the groupings in this study do not differ greatly from our current schemes of classification, but neither do they add much that was not already known about the relationships among these organisms.

What, then, can be said in favor of analyzing thermal transitions as

a means of determining nucleotide relationships? First, we are making very few assumptions with this technique. It is assumed that each time DNA is purified from cells of a given bacterial strain, the population of DNA fragments produced will be nearly identical. We assume that the fraction of the total hyperchromic change after each temperature increase is proportional to the amount of DNA denatured during that interval. Finally, we assume that similar organisms will have more stretches of nucleotide bases identical, and because of this, their DNA should melt in similar fashion. In short, we assume that thermal transition curves are reproducible, and that similar organisms should have transition curves that are more alike than those from organisms that are distantly related. With this minimal number of assumptions, there is nothing difficult to comprehend about the technique - it is simply a statistical analysis of data obtained when DNA is heat denatured. The technique is reproducible within limits determined by this study, and it is a technique that many laboratories are now performing routinely. Indeed, with a few mathematical manipulations, data already collected for determination of base composition could be used for a more complete analysis of the thermal transition. Although it is convenient to program a computer to do the calculations, one could perform the computations quite readily on a desk-top calculator.

Nonetheless, the assumptions made for this technique probably are oversimplified. We do not know exactly what happens when a DNA solution is denatured. The more popular theory is that there is partial denaturation of certain regions along numerous chromosome fragments (Inman, 1966 and 1967) as opposed to all-or-none denaturation of individual fragments

(Crothers *et al.*, 1965). Although it may make little difference here what DNA does physically as long as it denatures the same way every time, it would be reassuring to know exactly what happens, and perhaps to confirm our interpretations of the thermal transition curve. In addition, the technique does not take into account any differences in molecular weight of the genomes. It is conceivable that a bacterium with a smaller genome (*e.g.*, *Mycoplasma pneumoniae*: G + C = 39 %, molecular weight, 4.8×10^8; Bak *et al.*, 1969) could have a denaturation curve nearly identical to that of a bacterium with a much larger genome (*e.g.*, *P. vulgaris*: G + C = 39 %, molecular weight, 5×10^9; Falkow *et al.*, 1964). Finally, it is not known how sensitive the characteristics of thermal transition curves are to minor differences in DNA composition. For instance, many plasmids are known for the enteric bacteria in particular. What effect would such an extrachromosomal piece of DNA have on the transition curve? In this study a comparison was made between two cultures of *E. coli*, one strain that is lysogenic for phage lambda (ATCC 23740), and the same strain that had been "cured" of lambda (ATCC 23739). No significant difference could be detected between these two strains, either because the technique is not sensitive enough to detect such a small piece of added DNA or because it is able to detect such a piece only when the added fragment has a much different base composition.

Experimentally, one could try to resolve this question by transferring a F'lac$^+$ plasmid (G + C = 50 % and approximate molecular weight, 1×10^7) to *Proteus vulgaris* (G + C = 39 %) as described by Falkow *et al.* (1964). Thermal transitions could be performed on purified DNA from these strains to determine what effects the added plasmid might have. Alternatively,

some hypothetical calculations can be made. Keeping in mind that the melting curve of bacteriophage lambda is unusual (Yabuki, Fuke, and Wada, 1971; Davidson and Szybalski, 1971) and that lambda is only one-hundredth the size of the average enteric chromosome (Cooper and Helmstetter, 1968; Davidson and Szybalski, 1971), one might calculate the presumed effect of adding any number of lambda genomes to whatever transition curve he chooses. Such calculations were performed for *E. coli*, *S. marcescens* and *P. vulgaris*. When a single copy of lambda was added, there were some very slight changes in the statistics of the hypothetical transition curves. None of the changes approached the limits of significance, even in those strains with a large difference in mean percent G + C from lambda. This would be expected because the addition of a single copy of lambda was assumed to make a difference of only 0.002 optical density units {0.01 x (400 → 550)}. This also suggests that the presence of plasmids should not cause major errors, because plasmids usually do not exceed 1 - 5 % of the host genome (Rownd, 1969; Kasamatsu and Rownd, 1970; Kontomichalou, Mitani, and Clowes, 1970; Clowes, 1972). As the number of hypothetical copies of phage lambda was increased, there were further changes in the characteristics, and these were greater for the two strains which were most dissimilar in base composition from lambda. With some reservations concerning the approach of this hypothetical experiment, it can be concluded that the technique might be sensitive enough to detect additions as small as one-hundredth the molecular weight of the genome; however, these changes would be very slight.

Within the last ten years there have been numerous studies on nucleic

acid hybridizations in the *Enterobacteriaceae*, and one of the goals of this investigation was to learn whether thermal transition data could be related to hybridization data. It must be pointed out that although DNA hybridization studies are one of the best ways of determining relatedness between organisms, these studies are most dependent upon the experimental conditions of the hybridization and the assumption that the genomes of the organisms tested have the same molecular weight (Gillis, DeLey, and DeCleene, 1970; Brenner *et al.*, 1972a). Brenner's group (1969; 1972a; 1972b; 1972c; 1973; Crosa *et al.*, 1973) has conducted the most comprehensive study of relatedness in the *Enterobacteriaceae* using hybridizations. As expected, their hybridizations have shown divergence in the sequence of nucleotides in the enterics; however, these investigators do not discuss classifications because they do not want to place any taxonomic significance on their data as yet.

Brenner *et al.* (1972a) tested the assumption that the various strains of *E. coli* differed largely by single base changes and found that some strains were only 85 % related. The only two strains studied here that had also been included in Brenner's work were *E. coli* K-12 and B. They had been found to be 94 % similar by hybridization, but were nearly indistinguishable on the basis of their thermal transition characteristics, suggesting that the sensitivity of thermal transition analysis is less than that obtainable in hybridizations. On the other hand, some strains of *E. coli* in this study were different from each other in several characteristics, emphasizing that their DNA was not identical. In this same study Brenner showed that most of the enterics have diverged about the same

amount (approximately 50 % homology with DNA from *E. coli* K-12). There are two exceptions: species of *Escherichia* and *Shigella* are closely related (greater than 85 % homology) and *Proteus mirabilis* and *P. vulgaris* have diverged at least twice as much as the other enterics. Although the data from thermal transitions agree in general with conclusions based on hybridization data, there are some specific areas of disagreement. For example, Brenner *et al.* (1972a) found *P. morganii* to be distinct from *E. coli*, while in this study *P. morganii* grouped with the *Escherichieae*. As have other taxonomists (*e.g.*, Gardner and Kato, 1972; and as does this study) Brenner *et al.* (1972b) suggest that the erwinias are a diverse group of organisms, but several of their findings are in disagreement with mine. For example, *E. carotovora*, one of the soft-rot species, was not closely related to any of the other enterics by hybridization, but by thermal transition it groups with the *Escherichieae*. Hybridization also indicated that *E. amylovora* was more closely related to the *Escherichieae* than to the klebsiellas and enterobacters, another finding in apparent opposition to what was indicated by thermal transitions.

Perhaps some of these discrepancies can be accounted for by the condensation of groupings into two dimensions, as was mentioned earlier, or perhaps thermal transition analysis simply lacks sensitivity. But it is disappointingly apparent that the results of hybridizations by other investigators cannot be compared readily with these results of analysis by thermal transitons. The failure of many investigators to use taxonomically recognized reference cultures results in a lack of common strains between studies, and makes valid comparisons difficult or impossible. At best one

can only conclude that analysis by thermal transition seems to be hinting at the same kind of relatedness detected by hybridizations. It becomes obvious that a comprehensive comparison between hybridization data and thermal transitions, employing the same reference cultures, would be very much in order.

SUMMARY

Spectral data obtained from thermal transitions and spectral analyses were evaluated. Evidence was presented that although spectral analysis is poorly reproducible, the technique of thermal transition analysis can be used confidently to compare nucleotide relationships in bacteria. Characteristics of DNA from 35 taxonomically significant enteric strains and four miscellaneous strains were computed, and groupings were based on these characteristics. These groupings were compared with the results of numerical taxonomic studies, accepted taxonomic schemes, and hybridization studies. Although nothing new was suggested for the taxonomy of the enterics, most of the groupings suggested by thermal transitions alone do not differ from those obtained by other methods. It can be concluded that analysis by thermal transitions is a relatively easy technique that provides useful data on the DNA of microorganisms.

LITERATURE CITED

American Type Culture Collection. 1972. The American type culture collection catalogue of strains. 10th ed. Author, Rockville, Md. 312 pp.

Bak, A. L., F. T. Black, C. Christiansen, and E. A. Freundt. 1969. Genome size of mycoplasmal DNA. Nature (London) 224:1209-1210.

Bellett, A. J. D. 1967. Numerical classification of some viruses bacteria, and animals according to nearest-neighbour base sequence frequency. J. Mol. Biol. 27:107-112.

Brenner, D. J., and S. Falkow. 1971. Genetics of the *Enterobacteriaceae*. C. Molecular relationships among members of the *Enterobacteriaceae*. Advan. Genet. 16:81-118.

Brenner, D. J., G. R. Fanning, K. E. Johnson, R. V. Citarella, and S. Falkow. 1969. Polynucleotide sequence relationships among members of *Enterobacteriaceae*. J. Bacteriol. 98:637-650.

Brenner, D. J., G. R. Fanning, F. J. Skerman, and S. Falkow. 1972a. Polynucleotide sequence divergence among strains of *Escherichia coli* and closely related organisms. J. Bacteriol. 109:953-965.

Brenner, D. J., G. R. Fanning, and A. G. Steigerwalt. 1972b. Deoxyribonucleic acid relatedness among species of *Erwinia* and between *Erwinia* soecies and other enterobacteria. J. Bacteriol. 110:12-17.

Brenner, D. J., A. G. Steigerwalt, and G. R. Fanning. 1972c. Differentiation of *Enterobacter aerogenes* from klebsiellae by deoxyribonucleic acid reassociation. Int. J. Syst. Bacteriol. 22:193-200.

Brenner, D. J., A. G. Steigerwalt, G. R. Fanning, and G. V. Miklos. 1973. Deoxyribonucleic acid relatedness in *Pectobacterium* species. Abstr. Annu. Meeting Amer. Soc. Microbiol. 73:42. (Abstr.)

Britten, R. J., and D. E. Kohne. 1968. Repeated sequences in DNA. Science 161:529-540.

Clowes, R. C. 1972. Molecular structure of bacterial plasmids. Bacteriol. Rev. 36:361-405.

Cooper, S., and C. E. Helmstetter. 1968. Chromosomal replication and the division cycle of *Escherichia coli* B/r. J. Mol. Biol. 31:519-540.

Cowan, S. T. *In press*. Family I. *Enterobacteriaceae in* N. E. Gibbons and R. E. Buchanan, eds. Bergey's manual of determinative bacteriology. 8th ed. Williams and Wilkins, Baltimore, Md.

Crosa, J., D. J. Brenner, W. H. Ewing, and S. Falkow. 1973. Deoxyribonucleic acid relatedness in the tribe *Salmonelleae*. Abstr. Annu. Meeting Amer. Soc. Microbiol. 73:42. (Abstr.)

Crothers, D. M. 1968. Calculation of melting curves for DNA. Biopolymers 6:1391-1404.

Crothers, D. M., N. R. Kallenbach, and B. H. Zimm. 1965. The melting transition of low-molecular-weight DNA: theory and experiment. J. Mol. Biol. 11:802-820.

Davidson, N., and W. Szybalski. 1971. Physical and chemical characterization of lambda DNA. Pages 45-82 *in* A. D. Hershey, ed. The bacteriophage lambda. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

DeLey. J. 1969. Compositional nucleotide distribution and the theoretical prediction of homology in bacterial DNA. J. Theoret. Biol. 22:89-116.

DeLey, J. 1970. Reexamination of the association between melting point, buoyant density, and chemical base composition of deoxyribonucleic acid. J. Bacteriol. 101:738-754.

DeLey, J., and J. VanMuylem. 1963. Some applications of deoxyribonucleic acid base composition in bacterial taxonomy. Antonie van Leeuwenhoek, J. Microbiol. Serol. 29:344-358.

DeLey, J., H. Cattoir, and A. Reynaerts. 1970. The quantitative measurement of DNA hybridization from renaturation rates. Europe. J. Biochem. 12:133-142.

Dickerson, R. E. 1971. The structure of cytochrome c and the rates of molecular evolution. J. Mol. Evolut. 1:26-45.

Difco Laboratories. 1953. Difco manual of dehydrated culture media and reagents. 9th ed. Author, Detroit, Mich. 350 pp.

Edwards, P. R., and W. H. Ewing. 1962. Identification of *Enterobacteriaceae*. 2nd ed. Burgess Publishing Company, Minneapolis, Minn. 258 pp.

Falkow, S., J. A. Wohlhieter, R. V. Citarella, and L. S. Baron. 1964. Transfer of episomic elements to *Proteus*. I. Transfer of F-linked chromosomal determinants. J. Bacteriol. 87:209-219.

Felsenfeld, G. 1968. Ultraviolet spectral analysis of nucleic acids. Pages 247-253 *in* L. Grossman and K. Moldave, eds. Methods in enzymology. Vol. 12. Nucleic acids. Part B. Academic Press, New York.

Felsenfeld, G., and S. Z. Hirschman. 1965. A neighbor-interaction analysis of the hypochromism and spectra of DNA. J. Mol. Biol. 13:407-427.

Focht, D. D., and W. R. Lockhart. 1965. Numerical survey of some bacterial taxa. J. Bacteriol. 90:1314-1319.

Gardner, J. M., and C. I. Kado. 1972. Comparative base sequence homologies of the deoxyribonucleic acids of *Erwinia* species and other *Enterobacteriaceae*. Int. J. Syst. Bacteriol. 22:201-209.

Gibbons, N. E., and R. E. Buchanan, eds. *In press*. Bergey's manual of determinative bacteriology. 8th ed. Williams and Wilkins, Baltimore, Md.

Gillis, M., J. DeLey, and M. DeCleene. 1970. The determination of molecular weight of bacterial genome DNA from renaturation rates. Europe. J. Biochem. 12:143-153.

Hirschman, S. Z., and G. Felsenfeld. 1966. Determination of DNA composition and concentration by spectral analysis. J. Mol. Biol. 16:347-358.

Hirschman, S. Z., M. Gellert, S. Falkow, and G. Felsenfeld. 1967. Spectral analysis of the intramolecular heterogeneity of λ DNA. J. Mol. Biol. 28:469-477.

Holley, R. W., J. Apgar, G. A. Everett, J. T. Madison, M. Marquisse, S. H. Merrill, J. R. Penswick, and A. Zamir. 1965. Structure of a ribonucleic acid. Science 147:1462-1465.

Inman, R. B. 1966. A denaturation map of the λ phage DNA molecule determined by electron microscopy. J. Mol. Biol. 18:464-476.

Inman, R. B. 1967. Denaturation maps of the left and right sides of the lambda DNA molecule determined by electron microscopy. J. Mol. Biol. 28:103-116.

Kasamatsu, H., and R. Rownd. 1970. Replication of R-factors in *Proteus mirabilis*: replication under relaxed control. J. Mol. Biol. 51:473-489.

Knittel, M. D., C. H. Black, W. E. Sandine, and D. K. Fraser. 1968. Use of normal probability paper in determining thermal melting values of deoxyribonucleic acid. Can. J. Microbiol. 14:239-245.

Kontomichalou, P., M. Mitani, and R. C. Clowes. 1970. Circular R-factor molecules controlling penicillinase synthesis, replicating in *Escherichia coli* under either relaxed or stringent control. J. Bacteriol. 104:34-44.

Krieg, R. E. 1968. Characterization of the DNA of microorganisms. Unpublished Ph. D. thesis. Library, Iowa State University, Ames, Iowa.

Krieg, R. E., and W. R. Lockhart. 1966. Classification of enterobacteria based on overall similarity. J. Bacteriol. 92:1275-1280.

Krieg, R. E., and W. R. Lockhart. 1970. Analysis of the thermal transition curves of deoxyribonucleic acid from microorganisms. Can. J. Microbiol. 16:989-995.

Lockhart, W. R. 1967. Factors affecting reproducibility of numerical classifications. J. Bacteriol. 94:826-831.

Lockhart, W. R., and K. Koenig. 1965. Use of secondary data in numerical taxonomy of the genus *Erwinia*. J. Bacteriol. 90:1638-1644.

Lockhart, W. R., and J. Liston, eds. 1970. Methods for numerical taxonomy. American Society for Microbiology, Bethesda, Md. 62 pp.

Mandel, M., and J. Marmur. 1968. Use of ultraviolet absorbance-temperature profile for determining the guanine plus cytosine content of DNA. Pages 184-195 *in* L. Grossman and K. Moldave, eds. Methods in enzymology. Vol. 12. Nucleic acids. Part B. Academic Press, New York.

Marmur, J. 1961. A procedure for the isolation of deoxyribonucleic acid from micro-organisms. J. Mol. Biol. 3:208-218.

Marmur, J., and P. Doty. 1962. Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. J. Mol. Biol. 5:109-118.

McCarthy, B. J. 1967. Arrangement of base sequences in deoxyribonucleic acid. Bacteriol. Rev. 31:215-229.

Miyazawa, Y. and C. C. Thomas, Jr. 1965. Nucleotide composition of short segments of DNA molecules. J. Mol. Biol. 11:223-237.

Nandi, U. S., J. C. Wang, and N. Davidson. 1965. Separation of deoxyribonucleic acids by Hg (II) binding and Cs_2SO_4 density-gradient centrifugation. Biochem. 4:1687-1696.

Obel, J., and D. Freifelder. 1972. The termini of bacteriophage T7 DNA. J. Mol. Biol. 64:439-447.

Padmanabhan, R., and R. Wu. 1972. Nucleotide sequence analysis of DNA. IV. Complete nucleotide sequence of the left-hand cohesive end of coliphage 186 DNA. J. Mol. Biol. 65:447-467.

Prouty, J. L. 1970. The characteristics of deoxyribonucleic acid isolated from bacteria in balanced growth. Unpublished Ph. D. thesis. Library, Iowa State University, Ames, Iowa.

Rownd, R. 1969. Replication of a bacterial episome under relaxed control. J. Mol. Biol. 44:387-402.

Sanderson, K. E. 1972. Linkage map of *Salmonella typhimurium*, edition IV. Bacteriol. Rev. 36:558-586.

Schildkraut, C. L., and S. Lifson. 1965. Dependence of the melting temperature of DNA on salt concentration. Biopolymers 3:195-208.

Schildkraut, C. L., J. Marmur, and P. Doty. 1962. Determination of the base composition of deoxyribonucleic acid from its buoyant density in CsCl. J. Mol. Biol. 4:430-443.

Seidler, R. J., and M. Mandel. 1971. Quantitative aspects of deoxyribonucleic acid renaturation: base composition, state of chromosome replication, and polynucleotide homologies. J. Bacteriol. 106:608-614.

Snedecor, G. W., and W. G. Cochran. 1967. Statistical methods. 6th ed. Iowa State University Press, Ames, Iowa. 593 pp.

Szybalski, W. 1971. The lambda DNA map. Page 779 *in* A. D. Hershey, ed. The bacteriophage lambda. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.

Taylor, A. L., and C. D. Trotter. 1972. Linkage map of *Escherichia coli* strain K-12. Bacteriol. Rev. 36:504-524.

Ulitzur, S. 1972. Rapid determination of DNA base composition by ultraviolet spectroscopy. Biochim. Biophys. Acta 272:1-11.

Votavova, H., J. Sponar, and Z. Sormova. 1968. Comparison of heterogeneity of deoxyribonucleic acids of different origin by the method of hyperchromic spectra. Collection Czech. Chem. Commun. 33:4356-4368.

Watson, J. D., and F. H. C. Crick. 1953. Genetic implications of the structure of deoxyribonucleic acid. Nature (London) 171:964-967.

Weiner, R. M. 1970. Determination of intramolecular heterogeneity of microbial DNA by spectral analysis. Unpublished Ph. D. thesis. Library, Iowa State University, Ames, Iowa.

Wu, R., and A. D. Kaiser. 1967. Mapping the 5'-terminal nucleotides of the DNA of bacteriophage λ and related phages. Proc. Nat. Acad. Sci. U.S.A. 57:170-177.

Wu, R., and E. Taylor. 1971. Nucleotide sequence analysis of DNA. II. Complete nucleotide sequence of the cohesive ends of bacteriophage λ DNA. J. Mol. Biol. 57:491-511.

Yabuki, S., M. Fuke, and A. Wada. 1971. The fine structures in melting curves of deoxyribonucleic acids of bacteriophage lambda. J. Biochem. 69:191-207.

ACKNOWLEDGEMENTS

I would like to express my appreciation to my major professor, Dr. W. R. Lockhart, for his advice and assistance throughout this program; to the graduate faculty, especially Drs. J. G. Holt and P. A. Pattee, for their interest and guidance; to my fellow graduate students for their suggestions; and to my wife, Eileen, for her devotion and patience.